EEG segmentation through time-varying PCA

Maxime Rio, Axel Hutt, Bernard Girau

EEG segmentation through time-varying PCA

Maxime Rio, Axel Hutt and Bernard Girau
Loria, Campus Scientifique, BP329
54506 Vandoeuvre-lès-Nancy, France
maxime.rio@loria.fr

***Keywords*— Evoked related potentials, principal components analysis, hidden Markov model, variational inference**

I. Introduction

EEG recordings have been used as a tool to measure the brain activity for a long time. Especially, through the stimulus evoked experimental paradigm, numerous cortical activation patterns have been discovered, as sequences of evoked related potentials (ERP) components.

Our aim is to use tools from machine learning in order to systematically investigate ERP. More precisely, we focus on detecting cognitive components and localize them in time and space. The main issue is to deal with as few trials as possible.

II. Method

A. Current methods

Generally, ERP are assumed to be generated by a set of sources inside the brain. In order to extract these hidden sources, principal components analysis (PCA) and, more recently, independent components analysis (ICA) are two main approaches.

Both techniques project data on a lower dimensional space, representing the signals of hidden sources. Data records are a linear combination of these signals. The weights are given by a matrix which represents the contribution of each hidden source to each recorded channel. It may be considered as a kind of *activity map* w.r.t. the hidden sources.

Usually, these techniques are applied on a truncated dataset, within a specific time-window where ERP components are expected. It brings limitations since the time-window has to be known in advance. In addition, only one activity map and a fixed number of hidden sources are computed for the dataset.

B. Proposed probabilistic model

In order to overcome these limitations, we have developed a new technique relying on probabilistic inference. The goal is to extract the same kind of information as by conventional techniques, within several time windows, that are detected automatically. Thus, the activity map and the sources may change over time.

Our algorithm is based on a continuous hidden Markov model (HMM) while, in contrast to previous studies, here we propose to use probabilistic PCA analyzers instead of Gaussian components. The HMM segments time-series into time-windows. Each time-window is modeled by a probabilistic PCA analyzer.

This model has been set up within a Bayesian framework. Therefore, with appropriate prior probabilities, an automatic relevance determination (ARD) mechanism infers the number of PCA analyzers and, for each analyzer, the number of dimensions (i.e. the number of hidden sources) [1][2]. This approach tends to avoid overfitting. Moreover, The inference of the Bayesian model is conducted using the variational approximation [2].

III. Results

A. Synthetic data

A first experiment has been made on with 10-dimensional synthetic time-series. The model performed very well to isolate the different time-windows, i.e. the different states of the HMM.

B. BCI P300 dataset

The model was also used on a trial average of EEG records representing P300 evoked potentials [3].

As expected, the time-series was split into several parts. Interestingly, these parts were correlated to several known ERP components. Furthermore, the different activity maps extracted by each PCA analyzer correspond to possible activated areas on the scalp for the observed task.

IV. Conclusion

We propose a new probabilistic model combining two known approaches (HMM and PCA) in order to add the time information to the conventional PCA approach.

First results, on synthetic and real datasets, are promising. However, more tests have to be done on real datasets, in order to assess the correlation of inferred transitions and ERP components. Future work is also concentrated on the robustness of the model w.r.t. to the number of trials used.

V. Bibliography

[1] Bishop, C. Variational principal components. In Proceedings Ninth International Conference on Artificial Neural Networks, ICANN'99, Volume 1, pp. 509–514. IEE.

[2] Beal, M.J. Variational Algorithms for Approximate Bayesian Inference. PhD. Thesis, Gatsby Computational Neuroscience Unit, University College London, 2003

[3] Wolpaw, J. R., Schalk, G. and Krusiensk, D. Wadsworth. BCI Competition III, dataset II, “P300 speller paradigm”. 2004